



Partial Translation of JP-A 10-222606

Part A (Page 2)

[Claim 1]

An optical character reading apparatus comprising:

.
.
.
.

character entry frame detecting means for detecting a character entry frame position from a character entry frame and character image detected by the frame and character image detecting means;

character frame removing means for removing a character entry frame based on character entry frame position data detected by the character entry frame detecting means;

Part B (Page 3)

[0025] Then, at the character frame detecting part 3, vertical and horizontal projection histograms are generated from one line character image stored in the line buffer 1 (f3), and positions of the vertical frame line and horizontal frame line are detected (f4). With respect to judgement of the frame line positions, frame lines are judged as being present when values of the histograms exceed thresholds which are vertically and horizontally predetermined.

[0026] Figs. 3 and 4 are examples of vertical and horizontal

JEST AVAILABLE COPY

Our ref. : 61708/00R00384/US/JJF

projection histograms of Fig. 2, and positions denoted by reference characters a to h in Figs. 3 and 4 indicate positions where frame lines exist.

[0027] The character frame removing part 4 removes a character entry frame from frame line positions which are obtained from the projection histogram (f5).

OPTICAL CHARACTER READING DEVICE

Patent Number: JP10222606
Publication date: 1998-08-21
Inventor(s): MATSUYAMA YOSHIYUKI
Applicant(s):: MATSUSHITA ELECTRIC IND CO LTD
Requested Patent: ☐ JP10222606
Application Number: JP19970034266 19970204
Priority Number(s):
IPC Classification: G06K9/20
EC Classification:
Equivalents:

Abstract

PROBLEM TO BE SOLVED: To correctly segment a character which is described by being brought into contact with a character entry frame printed by a non drop-out color or being protruded so as to recognize it.

SOLUTION: A labeling processing part 2 adds numbers for every character group of one line character image. A character frame detecting part 3 detects a character frame position and a character frame removing part 4 removes a character frame. A character pattern lack correcting part 5 obtains a cut end point based on data where a part of a character removed by a character frame lacking part 4 is labeled by the labeling processing part 2 and corrects the lacking part of the character. By this means, segmenting is correctly recognized even at the time of the contact of the character frame with the character and protrusion.

Data supplied from the esp@cenet database - I2

(51) Int.Cl.⁶
G 0 6 K 9/20識別記号
3 4 0F I
G 0 6 K 9/20

3 4 0 L

審査請求 未請求 請求項の数 1 F D (全 5 頁)

(21) 出願番号 特願平9-34266

(22) 出願日 平成9年(1997) 2月4日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 松山 好幸

神奈川県横浜市港北区綱島東四丁目3番1

号 松下通信工業株式会社内

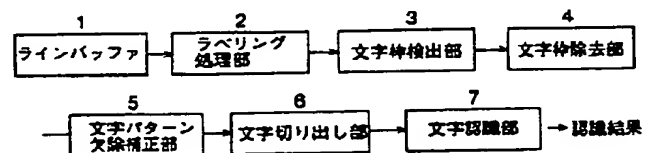
(74) 代理人 弁理士 青木 輝夫

(54) 【発明の名称】 光学的文字読取装置

(57) 【要約】

【課題】 非ドロップアウトカラーによって印刷された文字記入枠に接触したりはみ出したりして記載された文字を正確に切り出し認識することを可能とする。

【解決手段】 ラベリング処理部2は、1行分の文字イメージに対して、文字塊毎に番号を付加する。文字枠検出部3により文字枠位置を検出、文字枠除去部4により、文字枠を除去する。文字パターン欠除補正部5は、文字枠欠除部4によって除去された文字の一部をラベリング処理部2によってラベル付けされたデータを基に、切断された端点を求め、文字の欠除部分を補正する。この手段により、文字枠と文字が接触したり、はみ出した場合でも、正確に切り出し認識することができる。



【特許請求の範囲】

【請求項1】 帳票等の処理対象用紙に印刷された文字記入枠内に記録された文字を、前記処理対象用紙のシートイメージを検出し、同シートイメージから文字パターンを切り出して文字認識処理を行って、読取る光学的文字読取装置が、
前記処理対象用紙のシートイメージから前記文字記入枠及び文字のイメージを検出する枠・文字イメージ検出手段と、

この枠・文字イメージ検出手段によって検出された文字記入枠及び文字のイメージにラベリング処理を施して文字塊毎にラベル付けを施すラベリング処理手段と、

前記枠・文字イメージ検出手段によって検出された文字記入枠及び文字のイメージから文字記入枠位置を検出する文字記入枠検出手段と、

前記文字記入枠検出手段によって検出された文字記入枠位置データに基づいて文字記入枠を除去する文字枠除去手段と、

前記文字枠除去手段によって除去された文字の欠如部分を、前記ラベリング処理手段によってラベル付けされたデータを基に補正する文字パターン欠如補正手段と、
前記文字パターン欠如補正手段によって文字記入枠に重なった文字イメージの塊を検出して、文字パターンの検出を行う文字切り出し手段とを、

前記文字切り出し手段によって切り出された文字イメージと標準文字パターンとのマッチングを行って、文字を認識する文字認識手段とを、

有して構成することを特徴とする光学的文字読取装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、帳票等に記入された文字の読取りを行う光学的文字読取装置に関する。

【0002】

【従来の技術】一般に、帳票等の用紙に記入された文字を読取る光学的文字読取装置においては、帳票等の用紙に設けられた文字記入位置を示す枠（文字記入枠）内に記録された文字を切り出して認識している。

【0003】このために、文字記入枠が非ドロップアウトカラーによって印刷されている場合、文字は、文字記入枠と接触したり、はみ出さないことが、文字記入条件となっている。

【0004】しかし、現在は帳票と用紙への記録は、手書きによったり、スタンプによって多く行われ、このために、記録文字が、文字記入枠に接触したり、はみ出したりする場合が多く存在することとなる。

【0005】このような問題を解決するために、従来光学的文字読取装置として、例えば、特開平3-282895号公報に記載されたものがある。

【0006】これによれば、従来の光学的読取装置は、処理対象とする用紙に設けられた文字記入枠に記録され

た文字を、前記用紙のシートイメージを検出し、同シートイメージから文字パターンを切り出して文字認識処理を行うことによって読取る場合、前記シートイメージから前記文字記入枠の枠イメージを検出する枠イメージ検出手段と、前記枠イメージ検出手段によって検出された枠イメージから、前記文字パターンと重複しない除去すべき部分を検出する除去イメージ検出手段と、前記除去すべき部分を、前記シートイメージから除去するイメージ除去手段とを具備し、前記イメージ除去手段による除去の結果得られたシートイメージから、文字パターンの検出を行うように構成しており、このような構成によれば、文字記入枠と接触したり、はみ出して文字が記録されることによって、文字パターンと枠イメージとが重複する部分がある場合であっても、枠イメージの重複しない部分のみが除去され、不必要な枠イメージが除去されたシートイメージが切り出され、接触等すること無く記録された文字の文字パターンと同一の品質で文字認識処理を行うものであるとしている。

【0007】

【発明が解決しようとする課題】しかしながら、従来の光学的文字読取装置では、手書きによって記載された文字やスタンプによって押された文字など、文字記入枠に接触したり、はみ出したりする文字が存在した場合、はみ出した文字イメージから認識に不必要な文字記入枠を除去してしまうと、文字の一部が欠除し、文字認識結果に影響が出てしまうことがあった。

【0008】本発明は、非ドロップアウトカラーによって印刷された文字記入枠に接触したりはみ出したりして記載された文字を正確に切り出し認識することが可能な光学的文字読取装置を提供することを目的としている。

【0009】

【課題を解決するための手段】上記目的に鑑み、本発明による光学的文字読取装置は、枠・文字イメージ検出手段によって検出された文字記入枠及び文字のイメージにラベリング処理を施して文字塊毎にラベル付けを施すラベリング処理手段と、このラベリング処理手段によってラベル付けされたデータを基に、文字枠除去手段によって除去された文字の欠如部分を補正する文字パターン欠如補正手段と、この文字パターン欠如補正手段によって文字記入枠に重なった文字イメージの塊を検出して、文字パターンの検出を行う文字パターン切り出し手段とを有することを特徴とする。

【0010】これにより、非ドロップアウトカラーによって印刷された文字記入枠に接触したりはみ出したりして記載された文字を正確に切り出し認識することが可能となった。

【0011】

【発明の実施の形態】請求項1記載の光学的文字読取装置は、帳票等の処理対象用紙に印刷された文字記入枠内に記録された文字を、前記処理対象用紙のシートイメー

ジを検出し、同シートイメージから文字パターンを切り出して文字認識処理を行って、読取る光学的文字読取装置が、前記処理対象用紙のシートイメージから前記文字記入枠及び文字のイメージを検出する枠・文字イメージ検出手段と、この枠・文字イメージ検出手段によって検出された文字記入枠及び文字のイメージにラベリング処理を施して文字塊毎にラベル付けを施すラベリング処理手段と、前記枠・文字イメージ検出手段によって検出された文字記入枠及び文字のイメージから文字記入枠位置を検出する文字記入枠検出手段と、前記文字記入枠検出手段によって検出された文字記入枠位置データに基づいて文字記入枠を除去する文字枠除去手段と、前記文字枠除去手段によって除去された文字の欠如部分を、前記ラベリング処理手段によってラベル付けされたデータを基に補正する文字パターン欠如補正手段と、前記文字パターン欠如補正手段によって文字記入枠に重なった文字イメージの塊を検出して、文字パターンの検出を行う文字切り出し手段とを、前記文字切り出し手段によって切り出された文字イメージと標準文字パターンとのマッチングを行って、文字を認識する文字認識手段とを、有して構成することを特徴とするものである。

【0012】次に、本発明の実施の形態について、図面を参照しながら説明する。

【0013】図1は、本発明の実施の形態の主要部分の構成を示すブロック図である。処理対象となる帳票は光学的に走査することによって帳票イメージを得る。

【0014】1は、ラインバッファで、この帳票イメージから文字と文字記入枠が記載されている1行分のイメージを検出し、格納している枠・文字イメージ検出手段を構成する。

【0015】2は、ラベリング処理部で、ラインバッファ1に格納された1行分のイメージにラベリング処理を施して文字塊毎に番号を付けるラベリング処理手段を構成している。

【0016】3は、文字記入枠検出部で、ラインバッファ1に格納された1行分の文字イメージから縦・横方向の射影ヒストグラムをとり、文字枠位置を検出する文字記入枠検出手段を構成する。

【0017】4は、文字枠除去部で、文字枠検出部3によって検出された文字枠位置から文字枠を除去する文字枠除去手段を構成している。

【0018】5は、文字パターン欠如補正部で、文字枠除去部4によって除去された文字の一部をラベリング処理部2によってラベル付けされたデータを基に、文字の欠如部分を補正する文字パターン欠如補正手段を構成する。

【0019】6は、文字切り出し部で、文字パターン欠如補正部5で補正された文字イメージをラベル付けされた文字塊毎に切り出す文字切り出し手段を構成する。

【0020】7は、文字認識部で、文字切り出し部6に

より切り出された文字イメージと標準文字パターンとのマッチングが行われ、その結果が文字認識結果として出力する文字認識手段を構成している。

【0021】以上のように構成する実施の形態について、図2乃至図5を参照しながら、図6に記載されたフローチャートに基づいて、その動作を説明する。

【0022】図2は、帳票等の文字記入枠が存在する処理対象用紙に記載された1行分の文字イメージの例を示している。

【0023】この1行分の文字イメージが、ラインバッファ1に格納される(f1)。

【0024】次に、ラベリング処理部2によって、文字イメージは、黒画素の塊毎に、ラベル番号が付加される(f2)。これは、ある位置にある黒画素を考えたとき、これに接する上、下、右、左上、左下、右上、右下の黒画素にも同じラベル番号を付加し、文字イメージの塊を分類する操作である。

【0025】次に、文字枠検出部3では、ラインバッファ1に格納された1行分の文字イメージから縦・横方向の射影ヒストグラムをとり(f3)、縦枠線と横枠線の位置を検出する(f4)。この枠線の位置の判定は、ヒストグラム値が前もって縦・横方向それぞれ設定した閾値を越えた場合に、枠線が存在すると判定する。

【0026】図3、図4は、それぞれ図2の縦方向・横方向の射影ヒストグラムをとった一例であり、図中のa~hの位置が枠線の存在する位置である。

【0027】文字枠除去部4は、前記射影ヒストグラムより得られた枠線位置から文字記入枠を除去する(f5)。文字パターン欠如補正部5では、文字枠除去によって切除された文字イメージを補正する。

【0028】図5は、図2の一部を拡大して示すもので、文字パターン欠如補正部5によって文字枠除去後、切除された文字イメージを補正する動作をの一例を示す図である。領域8は縦枠線を示しており、文字枠除去4部により黒画素が白画素に変換されて、枠線が除去されている。画素の中に記されている数字は、ラベル番号を示している。

【0029】まず、あるラベル番号、例えば、ラベル番号1に注目して、縦枠線の1画素左の縦ラインを上から下に操作し(f6)、ラベル番号1が連続している部分を探索し、その始点と終点との組み合わせを求める(f7)。

【0030】この例では、始点と終点との組は、点1p1と点1p2、点1p3と点1p4、点1p5と点1p6である。また、この点のアドレス(横方向x、縦方向y)も記憶しておく。

【0031】同様にして、縦枠線の1画素右の縦ラインを上から下に走査し、始点と終点との組み点rp1と点rp2、点rp3とrp4点、点rp5と点rp6とを求める。

【0032】次に、左側の始点 $1p1$ 、 $1p3$ 、 $1p5$ のそれぞれについて、 y 座標の値が右側の始点 $rp1$ 、 $rp3$ 、 $rp5$ の y 座標の値のどれに最も近いかを求め、切断された点のペアを求める($f8$)。

【0033】この例では、始点 $1p1$ と $rp1$ 、 $1p3$ と $rp3$ 、 $1p5$ と $rp5$ が互いにペアとなり、終点の $1p2$ と $rp2$ 、 $1p4$ と $rp4$ 、 $1p6$ と $rp6$ とがペア値なる。

【0034】次に、このペア2点のアドレスから傾きを、整数値(少数定位以下は切り捨て) sw 求め($f9$)、縦枠線の1画素左にある始点または終点から出発して、1画素右の位置から傾き分上下に移動した位置を黒画素(ラベル番号1)にする。この黒画素にした位置から同様の操作を繰り返し、縦枠線の最右側ラインまで継続する。

【0035】図5の点 $1p1$ と点 $rp1$ のペアで考えると、傾きは4(画素)/4(画素)=1で、点 $w1$ 、 $w2$ 、 $w3$ が黒画素(ラベル番号1)となる。

【0036】同様に、点 $1p4$ と点 $rp2$ のペアについても行い、点 $w4$ 、 $w5$ 、 $w6$ が黒画素(ラベル番号1)になる。

【0037】この後、点 $w1$ と $w4$ との間、点 $w2$ と点 $w5$ の間、点 $w3$ と $w6$ の間も、黒画素(ラベル番号1)にする($f10$)。

【0038】この操作をすべての縦枠線左右の2点のペアについて行い、文字パターン欠除の補正が完了する。横枠線に対する処理も同様に実施可能である。

【0039】切り出し部6では、文字パターン欠除補正部5で出力されたイメージを基に、ラベル付けされた文字塊毎に切り出す($f11$)。図5の例では、ラベル番号1の塊を1文字として切り出す。この切り出された文字イメージは、標準パターン辞書に格納されている文字

の標準パターンとマッチングが行われ、その結果が文字認識結果として出力される($f12$)。

【0040】

【発明の効果】以上のように、本発明によれば、文字パターン欠除補正手段を設けることにより、非ドロップアウトカラーによって印刷された文字記入枠に接触したり、はみ出したりして記載された文字を正確に切り出し認識することが可能となる等、光学的文字読取装置として優れた効果を発揮することとなる。

【図面の簡単な説明】

【図1】本発明の実施の形態における主要部分の構成を示すブロック図。

【図2】本発明の実施の形態における帳票等の処理対象用紙に記載された1行分の文字イメージの例を示す図。

【図3】本発明の実施の形態における縦方向の射影ヒストグラムの例を示す図。

【図4】本発明の実施の形態における横方向の射影ヒストグラムの例を示す図。

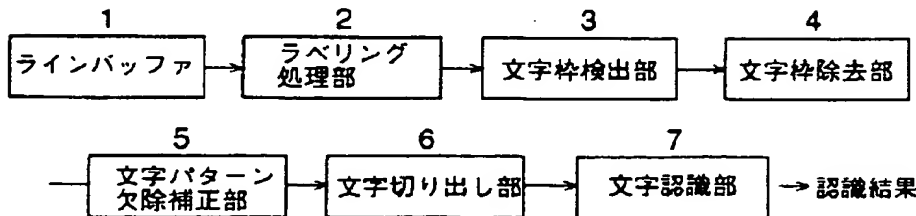
【図5】本発明の実施の形態における横方向の文字パターン欠除補正部の動作の説明図。

【図6】本発明の実施の形態における動作説明のためのフロー図。

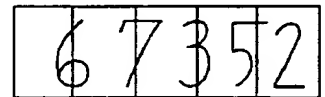
【符号の説明】

- 1 ラインバッファ(枠・文字イメージ検出手段)
- 2 ラベリング処理部(ラベリング手段)
- 3 文字枠検出部(文字枠検出部)
- 4 文字枠除去部(文字枠除去手段)
- 5 文字パターン欠除補正部(文字パターン欠除補正部)
- 6 文字切り出し部(文字切り出し手段)
- 7 文字認識部(文字認識手段)

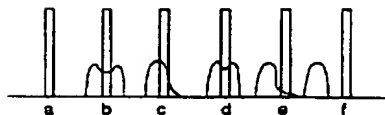
【図1】



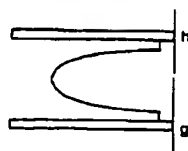
【図2】



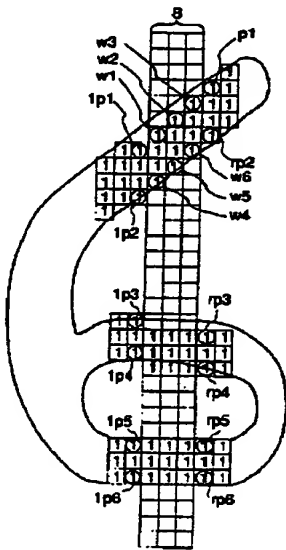
【図3】



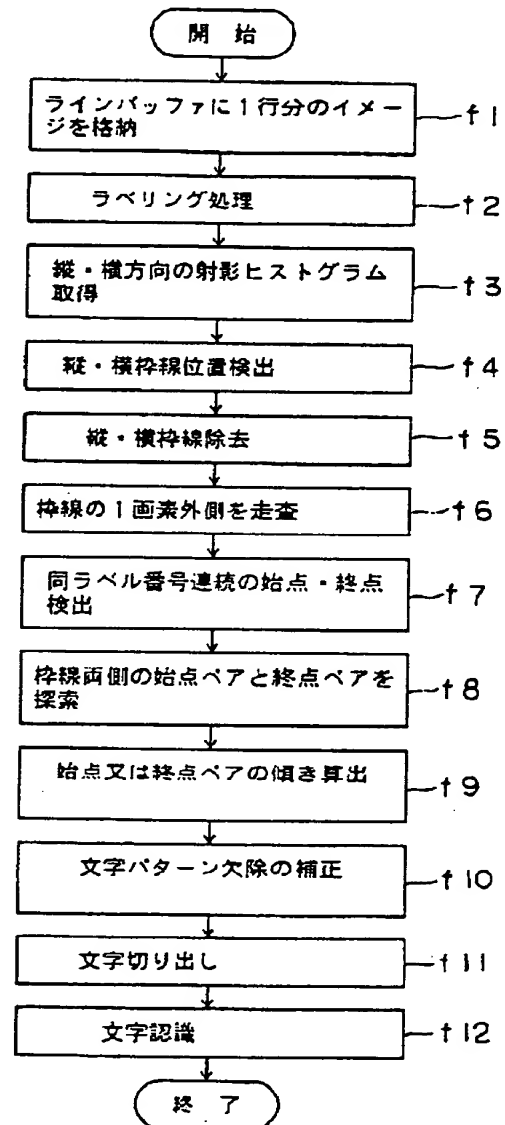
【図4】



【図5】



【図6】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.